© 1998 Oxford University Press
Nucleic Acids Research, 1998, Vol. 26, No. 12 **2955–2962**

Interaction between the N-terminal domain of human DNA topoisomerase I and the arginine–serine domain of its substrate determines phosphorylation of SF2/ASF splicing factor

Emmanuel Labourier, Ferdinand Rossi+, Imed-eddine Gallouzi, Eric Allemand, Gilles Divita[1] and Jamal Tazi*

Institut de Génétique Moléculaire de Montpellier (IGM), UMR 5535 CNRS, Université Montpellier II, CNRS - BP 5051, 1919, route de Mende, F34293 Montpellier Cedex 5, France and [1]Centre de Recherche de Biochimie Macromoléculaire, CNRS LP 9008–INSERM Unité 249, CNRS - BP 5051, 1919, route de Mende, F34293 Montpellier Cedex 5, France

Received February 24, 1998; Revised and Accepted April 27, 1998

ABSTRACT

Human DNA topoisomerase I, known for its DNA-relaxing activity, is possibly one of the kinases phosphorylating members of the SR protein family of splicing factors, *in vivo.* Little is known about the mechanism of action of this novel kinase. Using the prototypical SR protein SF2/ASF (SRp30a) as model substrate, we demonstrate that serine residues phosphorylated by topo I/kinase exclusively located within the most extended arginine–serine repeats of the SF2/ASF RS domain. Unlike other kinases such as cdc2 and SRPK1, which also phosphorylated serines at the RS domain, topo I/kinase required several SR dipeptide repeats. These repeats possibly contribute to a versatile structure in the RS domain thereby facilitating phosphorylation. Furthermore, far-western, fluorescence spectroscopy and kinase assays using the SF2/ASF mutants, demonstrated that kinase activity and binding were tightly coupled. Since the deletion of N-terminal 174 amino acids of Topo I destroys SF2/ASF binding and kinase activity but not ATP binding, we conclude that at least two distinct domains of Topo I are necessary for kinase activity: one in the C-terminal region contributing to the ATP binding site and the other one in the N-terminal region that allows binding of SF2/ASF.

INTRODUCTION

Several metazoan splicing factors involved in splice site selection are characterized by RNA recognition (RRM) and arginine–serine rich sequence motifs (RS domain) (1–3) that are essential for their function (4–6). These include members of SR proteins (SRp20, SF2/ASF or SRp30a, SC35 or SRp30b, SRp40, SRp55, SRp75) (1–3), U1 snRNP specific protein U1–70K (7,8) and the splicing factor U2AF (U2 snRNP auxilliary factor) which comprises two subunits (35K and 65K) (3,9,10). All these factors can mediate a network of protein interactions, through their RS domains, resulting in the stimulation and/or stabilization of complexes assembled at the 5′ or 3′ splice sites (11–15). Since the serines at the RS domains are phosphorylated (16–18), the protein phosphorylation could be a mechanism by which these protein–protein interactions might be regulated. This is supported by the fact that serine/threonine protein phosphorylation and dephosphorylation events play crucial roles in pre-mRNA splicing (17,19–21).

A major advance towards understanding the regulation of pre-mRNA splicing by protein phosphorylation was to isolate and characterize potential protein kinases and phosphatases that specifically target components of the splicing apparatus. Information regarding the possible kinases involved in the phosphorylation of SR proteins has recently emerged. The first evidence for a protein kinase activity with a specificity for the RS domain of SR proteins was presented by Woppmann *et al.* (22). The activity co-purifies with U1 snRNP and results in specific phosphorylation of both the C-terminal domain of the U1-snRNP 70K protein, which contains several arginine/serine rich clusters, and the RS domain of the splicing factor SF2/ASF. The kinase responsible for this activity remains unidentified, but the subset of sites in U1–70K phosphorylated by this kinase is the same as that detected *in vivo* (23). Recently, another SR protein kinase called SRPK1 was isolated and characterized (24,25). It is thought to regulate the intracellular localization of SC35 and other related splicing factors in the cell cycle (24). Purified SRPK1 can induce the disassembly of speckled intranuclear snRNP and SR protein structures in interphase nuclei. Since SR proteins are reported to be hyperphosphorylated in metaphase cells (24), SRPK1 may be the kinase that causes dynamic changes in the phosphorylation state of these structures during the cell cycle. However, SRPK1 is not the only protein kinase that mediates SR protein phosphorylation and redistribution in the cell. Colwill and his colleagues have provided evidence that ClK/Sty, a prototypical kinase with dual specificity,

*To whom correspondence should be addressed. Tel: +33 4 67 61 36 85; Fax: +33 4 67 04 02 45; Email: tazi@jones.igm.cnrs-mop.fr

+Present address: Columbia University, 630 West 168th Street, BB 20–22, New York, NY 10032, USA

capable of phosphorylating tyrosines as well as serines and threonines, is also involved in SR protein phosphorylation and localization (26).

We have shown that DNA topoisomerase I, which is a constitutively-expressed nuclear phospho-protein that localizes to active transcription sites (27–29), is a SR protein kinase (18). Consistent with this observation is the preferential inhibition of SR protein complete phosphorylation following treatment of HeLa cells with DNA topoisomerase I blockers (18). A more striking feature of this enzyme is that despite the absence of obvious sequence motifs homologous to known protein kinases (e.g. ATP binding site), it efficiently binds ATP. Photoaffinity labeling with 8-azidoadenosine-5′-triphosphate [α-^{32}P] combined with mutational analysis showed that the C-terminal was required for ATP binding (Rossi *et al.*, see accompanying paper). Here we report that the N-terminal 174 amino acids of human DNA topoisomerase I are required for the interaction with SF2/ASF protein *in vitro*. Furthermore, by utilizing a series of SF2/ASF truncation mutants, we also show that the RS domain is essential and sufficient for the interaction with Topo I. Furthermore, the phosphorylation mediated by topo I/kinase appears to have a specificity different from that of other SR kinases capable of phosphorylating the RS domain of SF2/ASF, implying that different mechanisms might be involved in the phosphorylation of this domain *in vivo*.

MATERIALS AND METHODS

Expression and purification of recombinant proteins

The wild type recombinant Topo I was produced following infection of monolayer cultures of Sf 21 cells with AcMNPV/hTOPI (a generous gift from A. M. Zhelkovsky, Tufts University, Boston), at a multiplicity of infection of 10. The virus encoding Topo-80 and Topo-δC were constructed as described (Rossi *et al.*, accompanying paper). All the methods for co-transfecting, isolating and maintaining recombinant baculovirus clones were performed according to Invitrogen manufacture protocols. With the exception of the N-terminally deleted Topo-70, starting with an engineered methionine immediately adjacent to residue Lys 175, which was generated and purified by L. Stewart from J.J. Champoux laboratory (30), recombinant Topo I mutant proteins were produced and purified following the same procedure as described by (18) and accompanying paper.

SF2/ASF was expressed in the TG1 bacteria strain transfected with plasmid containing ASF-1 cDNA, (a generous gift from J. Manley, Columbia University, New York). SF2/ASF deletion mutants at the RS domain were generated by PCR amplification of segments of ASF-1 cDNA spanning amino acid positions 1–197 (197C), 1–207 (207C), 1–215 (215C), 1–223 (223C), 1–236 (236C), 197–248 (197N), 206–248 (206N), 210–248 (210N), 216–248 (216N) and cloning these segments in pTrcHis A expression vector (Invitrogen). A large amount of these proteins was purified from inclusion bodies using procedure described by (31).

Kinase and topoisomerase assays

The reaction mixtures for protein kinase activity contained 100 ng of the recombinant Topo I proteins to be tested or equivalent kinase activity (determined by titration using SF2/ASF as standard substrate) of GST-SRPK1 (a gift from X.D. Fu) or p34 cdc2 (provided by M. Dorée laboratory), 300 ng of recombinant SF2/ASF mutants in buffer B (18) and 3 µCi [γ-^{32}P]ATP (3000 Ci/mmol) in 15 µl final volume and were incubated at 30°C for 30 min. The samples were then mixed with 5 µl of (3X) Laemmli loading buffer and applied to a 12% SDS–polyacrylamide gel (32).

To test the DNA relaxation activity of recombinant Topo I mutant proteins, 20 ng of each protein was incubated in buffer B with 500 ng supercoiled DNA plasmid in a final volume of 10 µl for 30 min at 30°C. The DNA plasmids in each sample were then extracted with phenol–chloroform, precipitated with ethanol and electrophoresed in a 0.8% agarose gel in TBE buffer. The DNA topoisomers were revealed by ethidium bromide staining.

Spectroscopic measurements

Fluorescence experiments were performed on a Fluorolog-II (Jobin Yvon) spectrofluorometer at 25°C. The proteins (25 nM) were incubated in a standard buffer containing 50 mM Tris–HCl, 2 mM $MgCl_2$ and 50 mM KCl. The binding of ATP, SF2/ASF or RS domain mutants was monitored by quenching of the intrinsic tryptophan fluorescence of recombinant Topo I proteins at 332 nm, upon excitation at 295 nm. The decrease in protein fluorescence was fitted to the appropriate form of the quadratic equation as already described (33,34). The best fit was obtained with the K_d corresponding to each protein and a maximal fluorescence quenching of 26% for recombinant proteins that bind SF2/ASF.

To avoid background fluorescence due to SF2/ASF itself, the single tryptophan residue from the SF2/ASF amino acid sequence was oxidized. To this end, 100 µg of SF2/ASF was incubated with 1 mM N-bromosuccinimide (NBS) freshly prepared in 50 mM Tris–acetate, pH 5 and 1 mM EDTA for 10 min, in the dark at 30°C. Unreacted NBS was neutralized by a 10-fold molar excess of free L-tryptophan and modified SF2/ASF was purified on nickel–agarose beads. Kinase reactions carried out with recombinant Topo I proved that this modified SF2/ASF behaved like untreated SF2/ASF (data not shown).

The circular dichroism (CD) spectra were recorded on a Mark V dichrograph (Jobin-Yvon, Paris) using 1 mm thick quartz cells and 500 µl of buffer [20 mM triethanolamine, 0.2 mM EDTA, 0.5 mM DTT, 42 mM $(NH_4)_2SO_4$, 70 mM NaF, 15% glycerol and 0.5 M guanidine] containing 85 µg of SF2/ASF, 197N or 197C proteins.

Far western analysis, 8N3-ATP cross-linking and chymotrypsin digestion of recombinant proteins

Recombinant SF2/ASF, 197C and full length Topo I (around 1 µg of each), or 1 µg of chymotrypsin-digested full length Topo I from each time point were run on 14% SDS–PAGE and transferred to nitrocellulose by electroblotting for 90 min in 10 mM CAPS, pH 11.0 transfer buffer, containing 10% methanol. To renature the proteins, the filters were treated as previously described (15) and probed with 10 µg of labeled SF2/ASF in 10 ml of binding buffer (15). To label SF2/ASF, 10 µg of purified recombinant protein was incubated with 800 U of Starfish purified cdc2 protein kinase in the presence of 20 µCi of [γ-^{32}P]ATP and 1 µM of cold ATP in buffer B for 1 h at 30°C. Unreacted nucleotides and cdc2 kinase were removed by binding labeled SF2/ASF to nickel–agarose beads and extensive washing of the beads with buffer B (18) before elution of labeled SF2/ASF with 1 M imidazole.

The protocol for photochemical cross-linking of 8N3-ATP to recombinant Topo-70 protein and digestion of labeled protein with chymotrypsin was described (Rossi *et al.*, accompanying paper).

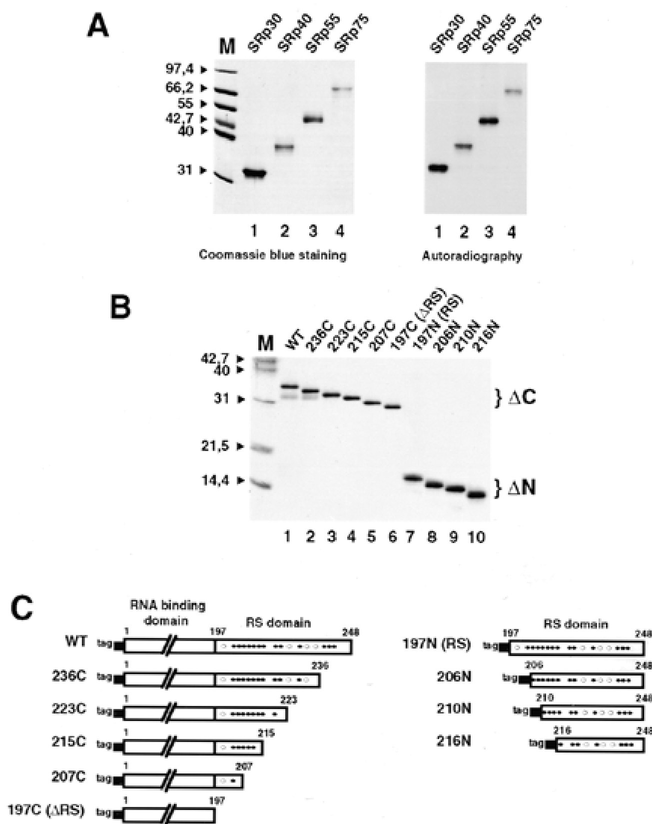


Figure 1. (**A**) Individual SR proteins purified from calf thymus (3) (left panel) were phosphorylated by purified topo I/kinase (right panel). Kinase assays were performed under the same conditions used to phosphorylate recombinant SF2/ASF protein. Coomassie blue staining (**B**) and schematic representation (**C**) of recombinant SF2/ASF mutants. Proteins were analyzed in 12% SDS–PAGE; M, molecular weight markers. In (C), open and closed circles represent SP and SR dipeptides, respectively.

RESULTS

The RS domain is a conserved feature of several members of the SR protein family. We tested whether well-characterized SR proteins can be utilized as substrate for the purified recombinant topo I/kinase, *in vitro*. Figure 1A shows that gel-purified individual SR proteins SRp30, SRp40, SRp55 and SRp75 derived from calf thymus (left panel) were all phosphorylated (right panel) by topo I/kinase, demonstrating that the enzyme utilized several proteins containing RS domain as substrate. For subsequent analysis we chose recombinant SF2/ASF (SRp30a) as a representative model substrate.

To examine the potential requirement of different domain(s) of SF2/ASF protein for the phosphorylation by topo I/kinase, we constructed several SF2/ASF deletion mutants missing various regions and domains. Mutants were constructed that lack either the entire N-terminal domain and therefore encode either the RS domain or portion of the RS domain, or only small regions of the RS domain (Fig. 1C). In addition, all the recombinant proteins were designed to contain a N-terminal His-tag to facilitate purification. They were expressed at high levels and purified from inclusion bodies of *Escherichia coli* by chromatography over an immobilized metal ion affinity column (His-bind resin). The electrophoretic analysis proved that the recombinant proteins were purified to homogeneity, since Coomassie blue-stained SDS gels revealed that the majority of the purified proteins were recovered as single predominant bands (Fig. 1B). Wild-type SF2/ASF (lane 1) and 236C mutant (lane 2) constituted exceptions as they were recovered as two bands. The largest forms of wild-type SF2/ASF and 236C had apparent molecular weights equivalent to those expected for the full-length proteins, whereas the smallest forms, although both contained the N-terminal his-tag, were generated by proteolysis at the C-terminus of the proteins.

Figure 2A shows that the RS domain alone confers the specificity of phosphorylation by topo I/kinase. No phosphorylation was detected with SF2/ASF that is deleted of its C-terminal half (197C, lane 2), while construct 197N, which lacks the entire N-terminal 196 residues, is phosphorylated to the same extent as wild type SF2/ASF (compare lanes 1 and 3). This demonstrated that topo I/kinase phosphorylation sites are confined to the C-terminal region and that the N-terminal half is not essential for phosphorylation. However, it is still possible that deletion of part(s) of SF2/ASF sequence perturbs the native structure of the protein. To test this possibility we performed CD analysis of full length SF2/ASF, 197C and 197N. The CD spectra for wild type and mutant SF2/ASF deleted of the RS domain were different (Fig. 2C), implying that deletion of the RS domain alters the structure of SF2/ASF. However, the CD spectrum of the isolated RS domain did not reveal strong optical activity at 207 and 222 nm, which is a very reliable predictor of α-helical content in the protein, suggesting that the RS domain has no significant α-helical content. It therefore appears that the RS domain has no well defined structure but rather a versatile one which is influenced by surounding sequences. As topo I/kinase displayed the same estimated K_m for wild type SF2/ASF and the RS domain alone (Fig. 2B), attention was turned to the possibility that an intact RS domain might be essential for phosphorylation by topo I/kinase.

The results of kinase assays performed with the various SF2/ASF mutants revealed that SF2/ASF constructs 215C, 223C and 236C which contain the entire SF2/ASF amino acid sequence but are truncated at amino acid positions 215, 223 and 236 respectively, were efficiently phosphorylated (Fig. 3, panel Topo I, lanes 2–4), showing that the COOH-terminal half of the RS region can be deleted without affecting the phosphorylation by topo I/kinase. In contrast, mutations that extend farther into the RS region (Fig. 3, panel Topo I, lanes 5 and 6) or where only the COOH-terminal half of the RS region was kept but not the N-terminal half (Fig. 3, lanes 9 and 10) were detrimental. Neither SF2/ASF truncated at amino acid position 207 (207C) nor constructs encoding the last 33 or 39 C-terminal residues (216N and 210N) were utilized as substrates by topo I/kinase. In contrast, the construct encoding the last 42 residues of the COOH-terminal region of SF2/ASF (206N) was efficiently phosphorylated. Therefore, we conclude that the region spanning residues 206–215 of SF2/ASF might contain an essential determinant for the phosphorylation by topo I/kinase, whereas the COOH-terminal half of the RS domain (217–248) is dispensable.

Further comparison of deletion mutants suggested that the number of repetitive SR dipeptides (206–215) is critical for the phosphorylation by topo I/kinase. Indeed, the mutants 215C and 206N, efficiently phosphorylated by topo I/kinase (Fig. 3, panel Topo I, lanes 4 and 8 respectively) have five and six repetitive SR dipeptides, respectively, but have different flanking sequences.

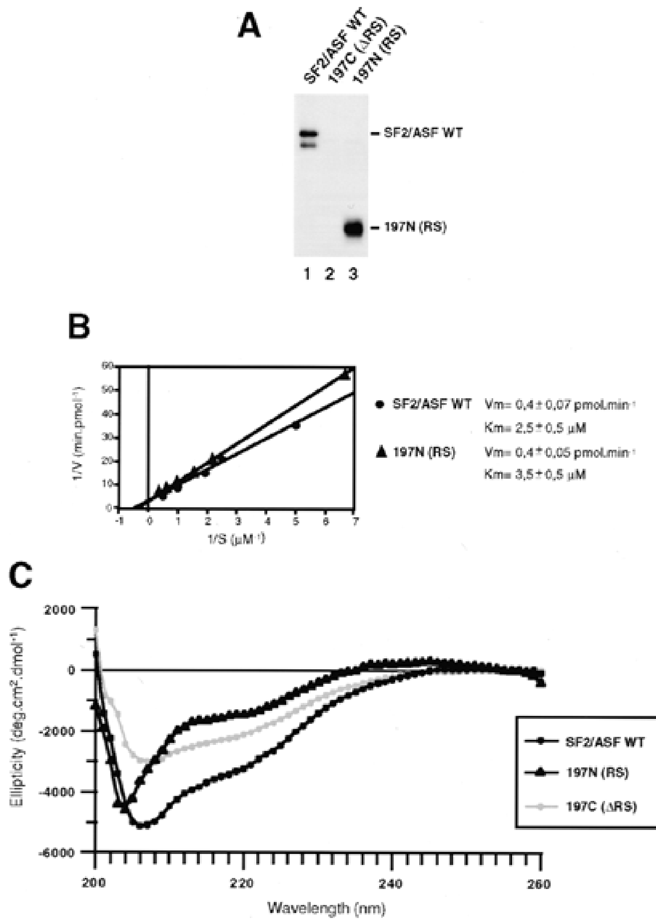


Figure 2. Topo I/kinase phosphorylates exclusively and efficiently the RS domain of SF2/ASF. (**A**) Kinase assays were performed as described in Materials and Methods using wild type SF2/ASF (lane 1), 197C (ΔRS) (lane 2) and 197N(RS) (lane 3) as substrates. Proteins were analyzed in 12% SDS–PAGE and detected by autoradiography. (**B**) To determine the K_m and V_m values, phosphorylation was kept linear with respect to reaction time (8 min) and topo I/kinase (100 ng). Kinase assays were carried out at 30°C in 15 µl buffer B with an increasing amount (3, 6, 11.25, 15, 22.5 and 30 pmol) of wild type SF2/ASF (closed circles) or (2.25, 9, 16.875, 22.5, 33.75 and 45 pmol) of 197N mutant (triangle) against one concentration of ATP (100 µM). The relative amount of phosphate transferred to radioactive substrate was quantitated by densitometry scanning of the gel using ImageQuant software version 3.22 and counting gel slices containing radioactivity excised from the Coomassie blue stained gel, for ^{32}P radioactivity in a liquid scintillation counter. A Lineweaver–Burk plot was derived from the reciprocal of kinase activities plotted against the reciprocal SF2/ASF or 197N mutant concentrations. (**C**) CD spectra of the indicated proteins are presented graphically as ellipticity values (deg.cm^2.mol^{-1}) values versus wavelength (nm).

However, the 210N deletion mutant, consisting of the same sequence as 206N but with only four repetitive SR dipeptides instead of six, was not phosphorylated by topo I/kinase (Fig. 3, panel Topo I, lane 9). This result suggests that it is sufficient to have five RS repeats and an additional flanking sequence at either end of SF2/ASF (N- or C-terminal) for maximal phosphorylation, using our assay conditions. Furthermore, substitutions of either Arg or Ser in the SR repeats (5) yielded mutants (RT, RG, GS and KS) that were not topo I/kinase substrates (data not shown). This demonstrates that threonine cannot replace serine as substrate for phosporylation by topo I/kinase and that phosphorylation of any of the serines requires the presence of arginine in the RS dipeptides.

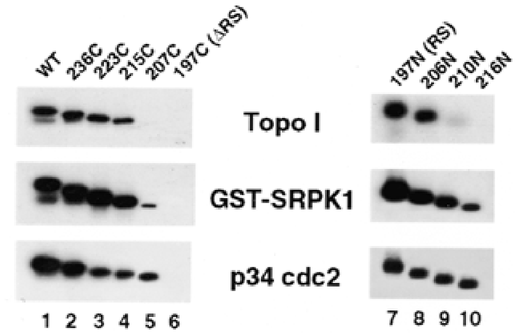


Figure 3. Phosphorylation of SF2/ASF mutants by topo I/kinase (panels Topo I), SRPK1 (panels SRPK 1) and p34 cdc2 kinase (panels p34 cdc2). Proteins were analyzed in 12% SDS–PAGE and detected by autoradiography. SF2/ASF mutants are indicated on the top of the panels.

Exclusive phosphorylation of the RS domain has also been observed using another SR protein kinase, SRPK1 (25). Like topo I/kinase, this enzyme fails to efficiently phosphorylate RT, RG, GS and KS mutants (35), suggesting that topo I/kinase and SRPK1 might have similar specificity. To test this possibility, we examined phosphorylation of all our SF2/ASF deletion mutants with recombinant GST-SRPK1. Figure 3 shows that although both enzymes phosphorylate several of the SF2/ASF mutants, there are several differences between SRPK1 (panel SRPK1) and topo I/kinase (panel Topo I). In particular, SRPK1 phosphorylates mutants 207C, 210N and 216N (panel SRPK1, lanes 5, 9 and 10) which are not phosphorylated by topo I/kinase (panel topo I, lanes 5, 9 and 10). The level of phosphorylation of SF2/ASF mutants by SRPK1 was dependent on the size of the RS domain, i.e. the number of serine residues. The smaller the RS domain length was, the weaker the phosphorylation by GST-SRPK1 (panel SRPK1, compare lanes 5 and 10 to lanes 1 and 7, respectively). In contrast, there was no obvious correlation between the levels of phosphorylation of the various substrates by topo I/kinase and the length of the RS domain, since all the SF2/ASF mutant substrates of topo I/kinase were roughly phosphorylated to the same extent. These results indicate that SRPK1 has a broader substrate specificity than topo I/kinase. They are also consistent with a previous report showing that SRPK1 was able to recognize several sites within full-length SF2/ASF RS domain (35).

The RS domain of SF2/ASF contains four serines followed by prolines at the P+1 position. These serines can be target sites for phosphorylation by cdc2 kinase which recognizes (S/T)-P consensus motif. We therefore compared the abilities of the different SF2/ASF mutants to be phosphorylated by this kinase. Wild type SF2/ASF was efficiently phosphorylated by cdc2 kinase (Fig. 3, panel cdc2, lane 1). As expected, the number of SP dipeptides determined the levels of phophorylation of each SF2/ASF mutant by cdc2 kinase (lanes 1–5, 7–10). No phosphorylation was observed with the 197C mutant which lacks all the SP dipeptides (lane 6). Mutants which have the same SP dipeptide contents showed the same phosphorylation levels (lanes 3–5 and 8–10, respectively). Wild type SF2/ASF (lane 1) and the 197N mutant (lane 7), which have the highest SP contents, had the highest level of phos-phorylation.

Thus, unlike SRPK1 and cdc2 kinases, which phosphorylate SF2/ASF mutants depending on small target sites (SR and SP

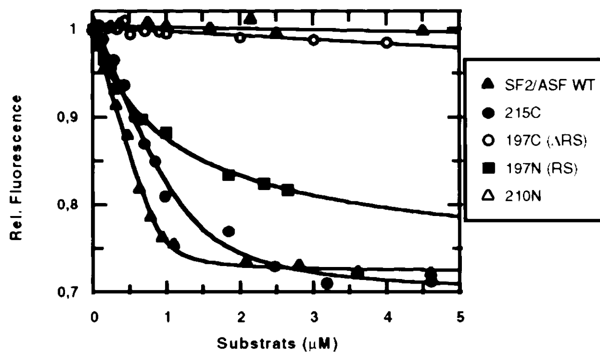


Figure 4. Binding of SF2/ASF mutants to Topo I revealed by tryptophan fluorescence quenching. Fluorescence experiments were performed as described in Materials and Methods using oxidized SF2/ASF mutants. The estimated K_d for the binding of each mutants to Topo I were as follows: full length SF2/ASF 0.56 μM, 215C 0.86 μM, 197N 0.89 μM and no quenching of Topo I fluorescence was observed with 197C and 210N.

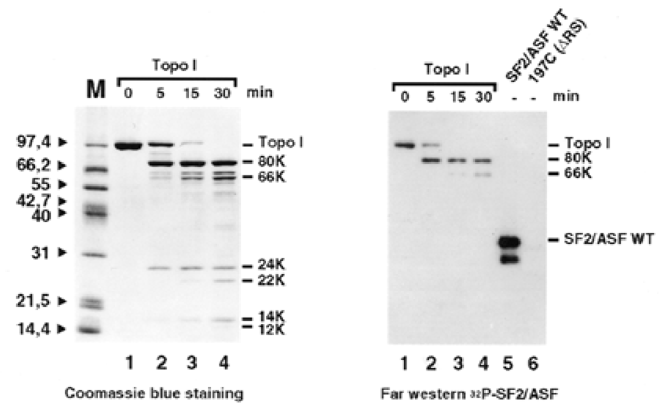


Figure 5. Far western analysis of purified Topo I digested with chymotrypsin. Protein samples corresponding to a time course digestion of wild type Topo I with chymotrypsin (Rossi *et al.*, accompanying paper) and recombinant SF2/ASF were transferred to nitrocellulose, renatured and probed with ^{32}P-labeled SF2/ASF (see Materials and Methods). Left panel shows Coomassie blue staining of chymotrypsin digested Topo I. Right panel autoradiography of far western analysis of chymotrypsin digested Topo I (lanes 1–4), wild type SF2/ASF (lane 5) and 197C (ΔRS) mutant (lane 6). The incubation times are indicated on the top and the relative size of proteolytic fragments, which were stained with labeled SF2/ASF, are indicated on the right of the panel.

dipeptides, respectively), topo I/kinase requires a minimum of five repetitive SR dipeptides for maximal phosphorylation. This Arg/Ser repeat may possibly contribute towards a versatile structure. Given that the RS domain is involved in protein–protein interaction (11–13,15) we asked whether specific phosphorylation of SF2/ASF depends on specific interaction(s) of this domain with defined domain (s) of topo I/kinase. To test this, the binding affinity of SF2/ASF to topo I/kinase was determined using tryptophane fluorescence. The single tryptophane residue contained in the SF2/ASF sequence was first oxidized in order to avoid intrinsic fluorescence due to SF2/ASF itself, while monitoring the quenching of the intrinsic fluorescence of recombinant topo I/kinase. This treatment, however, did not interfere with the ability of topo I/kinase to phosphorylate SF2/ASF as comparable amounts of phosphate were incorporated in untreated and treated SF2/ASF by topo I/kinase (data not shown). Figure 4 shows that increasing concentrations of oxidized SF2/ASF induces a quenching of 27% of Topo I fluorescence, indicating that this protein binds to topo I/kinase. The K_d value estimated from this analysis was 0.56 μM, which implies that SF2/ASF and topo I/kinase can form a stable complex. Similar analysis using SF2/ASF mutants strengthened the relationship between substrate binding and topo I/kinase-mediated phosphorylation. Mutants 215C and 197N which were efficiently phosphorylated by topo I/kinase, had similar binding to topo I/kinase as wild type SF2/ASF. In contrast, no binding was observed with mutants which were not utilized as substrate, i.e., 197C and 210N (Fig. 4). Note, however, that the quenching of Topo I fluorescence induced by 197N mutant (20%) was less than that caused by full length SF2/ASF. This difference could be easily explained by the fact that 197N mutant was smaller in size than full length SF2/ASF and therefore was most likely to cover a smaller area during interaction with Topo I.

To further confirm SF2/ASF–topo I/kinase interaction, we used far western blotting (11,12,15) which has been successfully used to show specific interactions between members of the SR protein family and other splicing factors. SF2/ASF immobilized on filter membranes and renatured was probed with ^{32}P-labeled SF2/ASF. The latter was labeled by phosphorylation of four SP dipeptides contained in its RS domain using cdc2 kinase and [γ-^{32}P]ATP. This phosphorylation did not interfere with the efficiency by which SF2/ASF RS domain is phosphorylated by topo I/kinase (data not shown), implying that the interaction between these two factors would not be affected by this phosphorylation. The specificity of binding of this modified SF2/ASF was confirmed by its ability to bind itself and U1–70 specific protein (data not shown), whereas it did not bind a recombinant SF2/ASF deleted of its RS domain (Fig. 5, right panel, compare lane 5 and lane 6). This is in agreement with previous studies showing that the interaction of SF2/ASF with itself was dependent upon an intact RS domain (11,15). Labeled SF2/ASF binds to topo I/kinase (Fig. 5, right panel, lane 1). The possibility that the ^{32}P-labeled probe was contaminated and/or binding to contaminating nucleic acids can be ruled out, because neither topo I nor SF2/ASF samples contained RNA or DNA (data not shown). The results all together, provide further evidence for an efficient interaction between SF2/ASF and topo I/kinase.

To identify the regions of topo I/kinase required for specific protein–protein interaction with SF2/ASF, chymotryptic fragments of the enzyme (Fig. 5, left panel, lanes 2–4) were analyzed by far western blotting (right panel, lanes 2–4). In addition to full length protein that was detected at the initial time of proteolysis (right panel, lane 2), two previously identified fragments of 80 and 66K (Rossi *et al.*, accompanying paper) appeared to bind SF2/ASF (lanes 2–4). However, no labeling of other proteolytic fragments was detected even after prolonged exposure of the autoradiogram, indicating that neither the 24K from the N-terminal nor the 14K from the C-terminal domains of topo I/kinase (Rossi *et al.*, accompanying paper) were recognized by SF2/ASF. Since the 80 and 66K fragments have the same N-terminal sequence but the 66K fragment lacks the C-terminal sequence, this suggested that SF2/ASF binds either near to the N-terminus of these two fragments or to the conserved core domain (residues Glu198–Ile651) of DNA topoisomerase I. We attempted to distinguish between these two possibilities by assessing the binding characteristics of a DNA topoisomerase I mutant lacking 174 N-terminal amino acids (Topo-70, Fig. 6A and B). Due to the absence of the rather extended NH_2-terminal domain, this protein

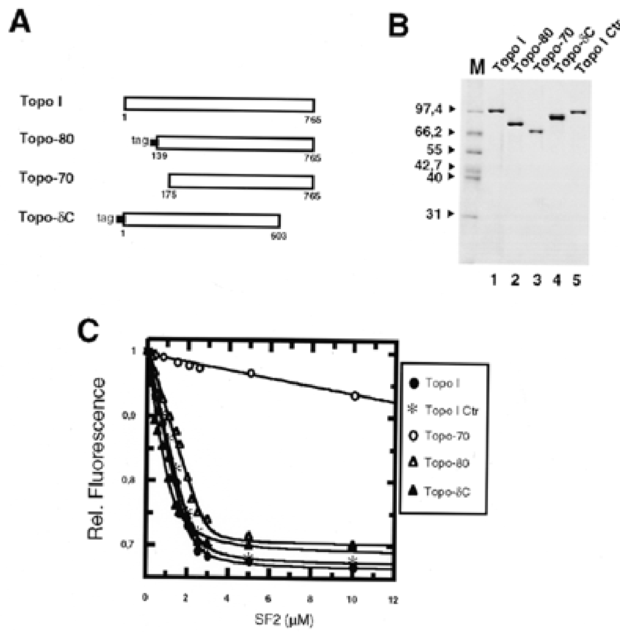


Figure 6. Schematic representation (**A**) and Coomassie blue staining of recombinant Topo I mutant proteins (**B**); quenching of the intrinsic fluorescence of each protein following binding of SF2/ASF (**C**). (A and B) Wild type (Topo I and Topo I Ctr), 80 kDa NH2-terminally truncated Topo I (Topo-80) with a histidine tag, 70 kDa NH2-terminally truncated Topo I (Topo-70), and COOH-terminally truncated Topo I (Topo-δC) with a histidine tag were purified after overexpression using baculovirus system (Rossi *et al.*, accompanying paper) and analyzed on a 10% SDS–PAGE (C) Oxidized SF2/ASF induced intrinsic fluorescence quenching of the recombinant proteins from which K_d values were estimated: Topo I 0.65 μM, Topo I Ctr 0.68 μM, Topo-80 0.92 μM and Topo-δC 0.58 μM. No significant quenching was observed with Topo-70. Topo I Ctr corresponds to full length topo I provided by Dr Champoux.

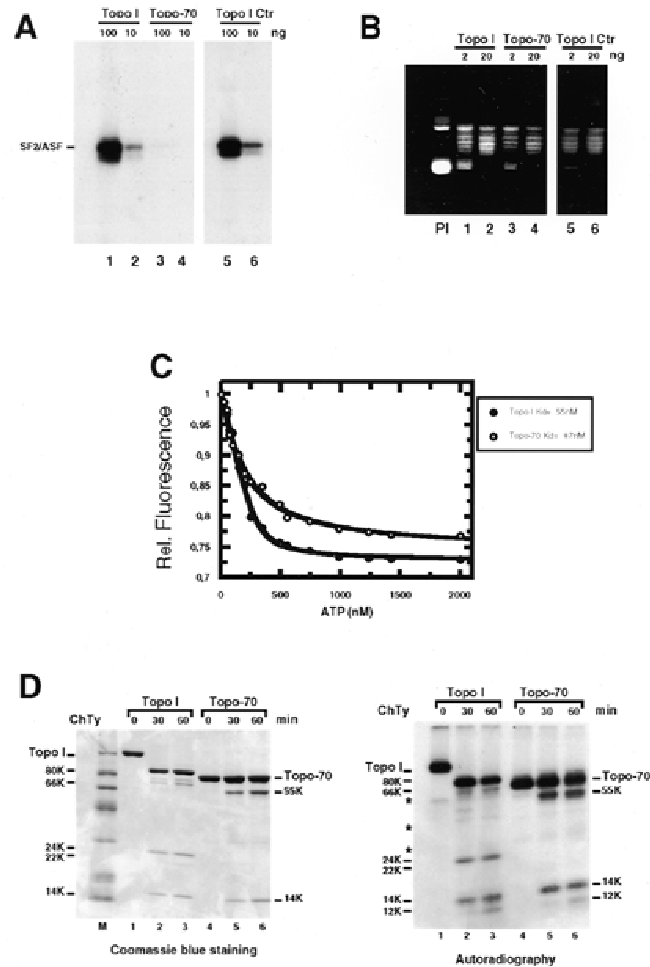


Figure 7. Recombinant Topo-70 has relaxing but not kinase activity. (**A**) Protein kinase assays were performed as described in Materials and Methods with 100 ng (lane 1) or 10 ng (lane 2) of wild type Topo I; 100 ng (lane 3), or 10 ng (lane 4) of purified Topo-70 from insect cells (35). Since Topo-70 was obtained from Dr Champoux's laboratory, we also performed kinase assays with 100 ng (lane 5) or 10 ng (lane 6) of wild type Topo I purified by the same laboratory. (**B**) DNA topoisomerase assays were as described in Materials and Methods with 2 ng (lane 1) or 20 ng (lane 2) of wild type Topo I; 2 ng (lane 3) or 20 ng (lane 4) of Topo-70; 2 ng (lane 5) or 20 ng (lane 6) of wild type Topo I from Dr Champoux's laboratory. (Pl) Supercoiled DNA incubated under the relaxation assay conditions in the presence of buffer only. (**C**) Determination of the K_d of ATP binding to Topo-70 (open circle) using the intrinsic fluorescence quenching of the protein and comparison with wild type Topo I (closed circles). (**D**) Wild type and Topo-70 were digested with chymotrypsin. Proteolytic fragments were analyzed on 10–16% SDS–polyacrylamide gel and revealed by Coomassie blue staining (left panel). Similar analysis was performed with wild type and Topo-70 cross-linked to [α^{32}P]8N3-ATP following irradiation with UV light (Rossi *et al.*, accompanying paper) and proteolytic fragments were detected by autoradiography (right panel). The incubation times are indicated on the top and the relative size of proteolytic fragments are indicated on the left of each panel. (*) correspond to proteolytic fragments resulting from exposure of recombinant proteins to UV light.

was previously shown to have a globular structure compared to wild type Topo I (36). Results showed that increasing the amount of oxidized SF2/ASF to 2 mM did not induce any quenching of tryptophan fluorescence of Topo-70, in contrast to the behavior of two other topo I/ kinase deletion mutants Topo-80 (deleted of 138 N-terminal residues) and Topo-δC (deleted of 162 residues from the C-terminal domain) in this assay (Rossi *et al.*, accompanying paper; Fig. 6C). Furthermore, immobilized Topo-70 did not interact with labeled SF2/ASF by far western analysis (data not shown). Hence, efficient binding of SF2/ASF to topo I/kinase requires amino acid sequences contained in the N-terminal region of the enzyme.

Interestingly, while it relaxed supercoiled DNA with the same efficiency as the full length Topo I (Fig. 7B), Topo-70 was inactive as a kinase (Fig. 7A). No phosphorylation of SF2/ASF was detected even when high concentrations of the enzyme were used (Fig. 7A, compare lanes 1 and 3). As expected, however, Topo-70 binds ATP with K_d only two times higher than wild type (Fig. 7C) and efficiently incorporates [α^{32}P]8N3-ATP (Fig. 7D). Digestion of photoaffinity labeled Topo-70 with chymotrypsin revealed distribution of the radioactivity between a 55K fragment from the N-terminus, the C-terminal 14K fragment and undigested Topo-70, confirming that even in the absence of the N-terminal region of topo I/kinase, the C-terminal domain has kept a conformation allowing it to bind ATP (Fig. 7D, compare lanes 1–3 and lanes 4–6). It is apparent from the size of these fragments (Fig. 7D, left and right panels) that chymotrypsin cleaved Topo-70 protein and wild type topo I at the same C-terminal site, suggesting that the N-terminal deletion of topo I/kinase does not affect the folding of its C-terminal part. Thus, the failure of Topo-70 to phosphorylate SF2/ASF appears to be due to its impaired ability to bind SF2/ASF.

DISCUSSION

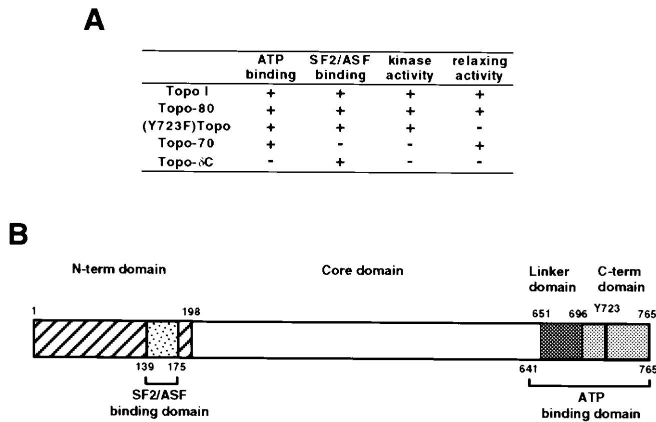
This study established that despite topo I/kinase, SRPK1 and cdc2 kinase phosphorylating serines in the RS domain of wild type SF2/ASF, these enzymes differ in their ability to phosphorylate mutants of SF2/ASF. Both SRPK1 and cdc2 kinase could phosphorylate serines preceding arginine or proline, respectively, in a wide variety of sequence backgrounds. In sharp contrast, there are limitations on the kind of sequences surrounding the serine that could be phosphorylated by topo I/kinase. This is probably related to the fact that a direct interaction between substrate and topo I/kinase is a prerequisite for the phosphorylation of SF2/ASF by this enzyme. Topo I mutant with impaired capacity to bind SF2/ASF (Topo-70) exerts no kinase activity on this protein. Similarly, no phosphorylation was observed with SF2/ASF mutants which do not interact with topo I/kinase.

Although our data do not address the mechanism governing the association of topo I/kinase with SF2/ASF, it can be proposed that charged residues contained within the interacting region of topo I/kinase (139–175) make electrostatic interactions with SF2/ASF. However, the 24K fragment generated following digestion of topo I/kinase with chymotrypsin failed to interact (by far western) with SF2/ASF, although it was present with the same stoichiometry as the 80K fragment. This fragment has a predicted PI value of 9.2 and would be expected to bind SF2/ASF if the charge alone was important for the interaction. Furthermore, SF2/ASF itself is a highly basic protein lacking sufficient numbers of acidic residues for stable electrostatic interactions (31,37). Thus charge alone does not appear to be the sole determinant for binding of SF2/ASF to topo I/kinase. Interestingly, inspection of the amino acid sequences contained in the N-terminal region of topo I/kinase indicates that while the sequence is highly conserved amongst higher eukaryotes, it is only partially conserved in yeasts (38). It is therefore possible that the amino acids contributing to the interaction with SF2/ASF and thereby to the kinase activity have been selected during the evolution of higher eukaryotes. Consistent with this view, yeast DNA topoisomerase I cannot interact with or phosphorylate SF2/ASF *in vitro*, unlike DNA topoisomerase I extracted from mouse, plants or insect cells (F.Rossi, unpublished observations). Topo I/kinase phosphorylates at least six members of the SR protein family on serines contained in the RS domain of these proteins, suggesting that conserved features are shared by all these proteins. The fact that SR proteins isolated from different multicellular organisms are recognized by an antibody, mAb104, that stains lateral loops on amphibian lampbrush and puffs on *Drosophila* polytene chromosomes (16,39,40), implies a high degree of conservation of SR proteins. To our knowledge this antibody fails to stain yeast proteins, implying that some of the features present in the SR proteins are missing in yeasts.

The finding that the residues spanning 139–175 are required for kinase activity represents another feature that distinguishes the kinase activity of topo I/kinase from DNA dependent topoisomerase I function. The N-terminal domain appears to be dispensable for DNA topoisomerase I activity *in vitro*. Deletion of this domain by spontaneous proteolysis (41) or by induced mutagenesis of the wild type protein (42) does not impede the relaxation activity. This domain, however, is essential for function *in vivo* (43). Overexpression of the wild type human DNA topoisomerase I resulted in a lethal phenotype in yeasts, whilst overexpression of a mutant lacking the N-terminal domain did not affect yeasts viability (43). Further experiments showed that this was due to accumulation of the protein in the cytoplasm of the yeast cell, indicating that the deleted domain is necessary for nuclear localization of DNA topoisomerase I. Consistent with this finding, the nuclear localization signal (NLS) from the SV40 large T antigen can substitute for the N-terminal domain of DNA topoisomerase I (141–210), restoring its nuclear localization and resulting in lethality in yeast (43). Since the sequences in this region do not conform to a classical NLS consensus sequence, it is possible that binding of SR proteins or other nuclear factors to DNA topoisomerase I may be a mechanism through which DNA topoisomerase I enters the nucleus. In this context, it may be relevant that the core, linker and C-terminal domain fold into a globular structure, while the N-terminal domain is highly extended (35) and thereby accessible for specific interaction with other proteins.

How might topo I/kinase contribute to the function of SR proteins? Human DNA topoisomerase I is a constitutively expressed nuclear phospho-protein that localizes to active transcription sites (27,28). Thus its kinase activity might allow this protein to participate in the coordination between transcription and splicing. Interestingly, DNA topoisomerase I blockers lead to the preferential inhibition of SR protein complete phosphorylation in HeLa cells (18). Furthermore, topo I/kinase associates with transcription factors, such as RNA polymerase I (44), nucleolin (45), TATA-binding protein (46,47) and p53 protein (48). It is worth noting that topo I relaxation activity was shown to be dispensable for both repression and activation of transcription in reconstituted transcription reactions (47,49). The effect of DNA topoisomerase I, in this case, was specific for TATA box-containing promoters and was mediated by the TATA-binding protein, suggesting a novel function for the enzyme, possibly as a kinase, in the regulation of transcription initiation by RNA polymerase II. Furthermore, sites of transcription in the nucleus colocalize with sites of splicing (50–52), implying that both processes are coupled to one another. An attractive model would be that topo I/kinase recruits SR proteins to the transcription initiation complex, and during the course of transcription these SR proteins are loaded onto the nascent transcript. The finding that topo I/kinase needs to associate with SR proteins to phosphorylate them fits nicely with this model.

Figure 8 shows a model depicting our interpretation of the results. The regions of human DNA topoisomerase I contributing to topo I/kinase were localized in the 765-amino acid protein predicted by D'Arpa *et al.* (53). Based on amino acid comparisons with other eucaryotic DNA topoisomerase I enzymes and susceptibility of different regions of the protein to cleavage with proteases, the model shows four distinct domains [adapted from Stewart *et al.* (42)]: the N-terminal domain (residues 1–198) is mostly unconserved except for small areas that are either part of topo I/kinase domains allowing binding of SF2/ASF or thought to be involved in nuclear localization of human DNA topoisomerase I; the core domain (residues 198–651, 54K) which is very well conserved and shown to be resistant to proteolysis; the linker domain (651–696), which is also unconserved; and the C-terminal domain (651–765), which contains in addition to the active tyrosine at the position 723, amino acid residues which are involved in the binding of ATP (Rossi *et al.*, accompanying paper). However, we are uncertain as to whether the SF2/ASF and ATP binding sites are confined to the regions established by deletion analysis of the wild type Topo I described here or extend to flanking sequences.



	ATP binding	SF2/ASF binding	kinase activity	relaxing activity
Topo I	+	+	+	+
Topo-80	+	+	+	+
(Y723F)Topo	+	+	+	-
Topo-70	+	-	-	+
Topo-δC	-	+	-	-

Figure 8. Human DNA topoisomerase I structural and functional domains. (**A**) Summary of the effect of Topo I mutations on ATP and SF2/ASF binding, and on relaxing and kinase activities of topo I/kinase. Absence or presence of binding and activity are indicated by – or + signs, respectively. (**B**) The 765-amino acid protein predicted by cDNA isolated by D'Arpa *et al.* (53) was divided into four structural domains [adapted from Stewart *et al.* (30)]: the unconserved N-terminal domain (residues 1–197), the conserved core domain (residues 198–651), the unconserved linker domain (residues 652–696) and the conserved C-terminal domain which contain the active tyrosine at position 723 (residues 697–765). Regions of DNA topoisomerase I contributing to topo I/kinase activity which correspond to SF2/ASF binding domain (residues 139–175) and ATP binding domain (residues 641–765) respectively are indicated below.

ACKNOWLEDGEMENTS

We are grateful to X. D. Fu and J. J. Champoux for providing recombinant GST-SRPK1 and Topo-70 or wild type topo I, respectively, as well as A. Krainer for providing RG, RT, GS and KS SF2/ASF mutant constructs. This work was supported by the Centre National de la Recherche Scientifique, The Ministère de l'Enseignement Supérieur et de la Recherche, the Association pour la Recherche contre le Cancer and the Fondation pour la Recherche Médicale.

REFERENCES

1 Fu,X.D. (1995) *RNA*, **1**, 663–680.
2 Manley,J.L. and Tacke,R. (1996) *Genes Dev.*, **10**, 1569–1579.
3 Zahler,A.M., Lane,W.S., Stolk,J.A. and Roth,M.B. (1992) *Genes Dev.*, **6**, 837–847.
4 Zuo,P. and Manley,J.L. (1993) *EMBO J.*, **12**, 4727–4737.
5 Càceres,J.F. and Krainer,A.R. (1993) *EMBO J.*, **12**, 4715–4726.
6 Zamore,P.D., Patton,J.G. and Green,M.R. (1992) *Nature*, **335**, 609–614.
7 Theissen,H., Etzerodt,M., Reuter,R., Schneider,C., Lottspeich,F., Argos,P., Lührmann,R. and Philipson,L. (1986) *EMBO J.*, **5**, 3209–3217.
8 Query,C.C., Bentley,R.C. and Keene,J.D. (1989) *Cell*, **57**, 89–101.
9 Zamore,P.D. and Green,M.R. (1989) *Proc. Natl Acad. Sci. USA*, **86**, 9243–9247.
10 Zamore,P.D. and Green,M.R. (1991) *EMBO J.*, **10**, 207–214.
11 Wu,J.Y. and Maniatis,T. (1993) *Cell*, **75**, 1061–1070.
12 Amerein,H., Hedley,M.L. and Maniatis,T. (1994) *Cell*, **76**, 735–746.
13 Zuo,P. and Maniatis,T. (1996) *Genes Dev.*, **10**, 1356–1368.
14 Reed,R. (1996) *Curr. Opin. Gen. Dev.*, **6**, 215–220.
15 Kohtz,J.D., Jamison,S.F., Will,C.L., Zuo,P., Lührmann,R., Garcia-Blanco,M.A. and Manley,J.L. (1994) *Nature*, **368**, 119–124.
16 Roth,M.B., Zahler,A.M. and Stolk,J.A. (1991) *J. Cell Biol.*, **115**, 587–591.
17 Tazi,J., Kornstädt,U., Rossi,F., Jeanteur,P., Cathala,G., Brunel,C. and Lührmann,R. (1993) *Nature*, **363**, 283–285.
18 Rossi,F., Labourier,E., Forné,T., Divita,G., Derancourt,J., Riou,J.F., Antoine,E., Cathala,G., Brunel,C. and Tazi,J. (1996) *Nature*, **381**, 80–82.
19 Tazi,J., Daugeron,M.C., Cathala,G., Brunel,C. and Jeanteur,P. (1992) *J. Biol. Chem.*, **267**, 4322–4326.
20 Xiao,S.H. and Manley,J.L. (1997) *Genes Dev.*, **11**, 334–344.
21 Parker,A.R. and Steitz,J.A. (1997) *RNA*, **3**, 1301–1312.
22 Woppmann,A., Will,C.L., Konstädt,U., Zuo,P., Manley,J.M. and Lührmann,R. (1993) *Nucleic Acids Res.*, **21**, 2815–2822.
23 Woppmann,A., Patschinsky,T., Bringmann,P., Godt,F. and Lührmann,R. (1990) *Nucleic Acids Res.*, **18**, 4427–4438.
24 Gui,J.F., Lane,W.S. and Fu,X.D. (1994) *Nature*, **369**, 678–682.
25 Gui,J.-F., Tronchère,H., Chandler,S.D. and Fu,X.-D. (1994) *Proc. Natl Acad. Sci. USA*, **91**, 10824–10828.
26 Colwill,K., Pawson,T., Andrews,B., Prasad,J., Manley,J., Bell,J.C. and Duncan,P.I. (1996) *EMBO J.*, **15**, 265–275.
27 Fleischmann,G., Pflugfelder,G., Steiner,E.K., Javaherian,K., Howard,G.C., Wang,J.C. and Elgin,S.C.R. (1984) *Proc. Natl Acad. Sci. USA,* **81**, 6958–6962.
28 Gilmour,D.S., Pflugfelder,G., Wang,J.C. and Lis,J.T. (1986) *Cell,* **44**, 401–407.
29 Heck,M.M., Hittelman,W.N. and Earnshaw,W.C. (1988) *Proc. Natl Acad. Sci.USA*, **85**, 1086–1090.
30 Stewart,L., Ireton,G.C. and Champoux,J.J. (1996) *J. Biol. Chem.*, **271**, 7602–7608.
31 Ge,H., Zuo,P. and Manley,J.L. (1991) *Cell*, **66**, 373–382.
32 Laemmli,U.K. (1970) *Nature*, **227**, 680–685.
33 Divita,G., Goody,R.S., Gautheron,D.C. and Di Pietro,A. (1993) *J. Biol. Chem.*, **268**, 13178–13186.
34 Rittinger,K., Divita,G. and Goody,R.S. (1995) *Proc. Natl Acad. Sci. USA*, **92**, 8046–8049.
35 Colwill,K., Feng,L.L., Yeakley,J.M., Gish,G.D., Caceres,J.F., Pawson,T. and Fu,X.D. (1996) *J. Biol. Chem.*, **271**, 24569–24575.
36 Stewart,L., Ireton,G.C., Parker,L.H., Madden,K.R. and Champoux,J.J. (1996) *J. Biol. Chem.*, **271**, 7593–7601.
37 Krainer,R.K., Mayeda,A. and Kozak,D. (1991) *Cell*, **66**, 383–394.
38 Gupta,M., Fujimori,A. and Pommier,Y. (1995) *Biochim. Biophys. Acta*, **1262**, 1–14.
39 Roth,M.B. and Gall,J.G. (1987) *J. Cell Biol.*, **105**, 1047–1054.
40 Roth,M.B., Murphy,C. and Gall,J.G. (1990) *J. Cell Biol.*, **111**, 2217–2223.
41 Liu,L.F. and Miller,K.G. (1981) *Proc. Natl Acad. Sci. USA*, **78**, 3487–3491.
42 Stewart,L., Ireton,G.C. and Champoux,J.J. (1996) *J. Biol. Chem.*, **271**, 7602–7608.
43 Alsner,J., Svejstrup,J.Q., Kjeldsen,E., Sorensen,B.S. and Westergaard,O. (1992) *J. Biol. Chem.*, **267**, 12408–12411.
44 Rose,K.M., Szopa,J., Han,F.S., Cheng,Y.C., Richter,A. and Scheer,U. (1988) *Chromosoma*, **96**, 411–416.
45 Bharti,A.K., Olson,M.O.J., Kufe,D.W. and Rubin,E.H. (1996) *J. Biol. Chem.*, **271**, 1993–1996.
46 Kretzschmar,M., Meisterernst,M. and Roeder,R.G. (1993) *Proc. Natl Acad. Sci. USA*, **90**, 11508–11512.
47 Merino,A., Madden,K.R., Lane,W.S., Champoux,J.J. and Reinberg,D. (1993) *Nature*, **365**, 227–232.
48 Gobert,C., Bracco,L., Rossi,F., Olivier,M., Tazi,J., Lavelle,F., Larsen,A.K. and Riou,J.F. (1996) *Biochemistry*, **35**, 5778–5786.
49 Shykind,B.M., Kim,J., Stewart,L., Champoux,J.J. and Sharp,P.A. (1997) *Genes Dev.*, **11**, 397–407.
50 Wuarin,J. and Schibler,U. (1994) *Mol. Cell. Biol.*, **14**, 7219–7225.
51 Spector,D.L. (1993) *Ann. Rev. Cell Biol.*, **9**, 265–315.
52 Jiménez-Garcia,L.F. and Spector,D.L. (1993) *Cell*, **73**, 47–59.
53 D'Arpa,P., Machlin,P.S., Ratrie,H., Rothfield,N.F., Cleveland,D.W. and Earnshaw,W.C. (1988) *Proc. Natl. Acad. Sci. USA,* **85**, 2543–2547.